PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C07K 14/47, 16/18, C12N 15/12, 15/00, 5/10, A61K 38/16, A01K 67/027

A1

(11) International Publication Number: WO 99/23252

(43) International Publication Date: 14 May 1999 (14.05.99)

(21) International Application Number: PCT/GB98/03375

(22) International Filing Date: 5 November 1998 (05.11.98)

(30) Priority Data:
9723412.4 5 November 1997 (05.11.97) GB
9818118.3 19 August 1998 (19.08.98) GB

(71) Applicant (*for all designated States except US*): ISIS INNOVATION LIMITED [GB/GB]; 2 South Parks Road, Oxford, Oxfordshire OX1 3UB (GB).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): BAKKENIST, Christopher, James [GB/GB]; 11 Woolaston Drive, Alsager, Stoke-on-Trent, Staffordshire ST7 2PL (GB). KORETH, John [IN/IN]; A-10/45 Mall Road, 110054 Delhi (IN). McGEE, James, O'Donnell [GB/GB]; 17 Cadagan Park, Woodstock, Oxfordshire OX40 1UW (GB).

(74) Agents: PRIVETT, Kathryn, L. et al.; Stevens, Hewlett & Perkins, 1 Serjeants' Inn, Fleet Street, London EC4Y 1LL (GB).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CANCER GENE

(57) Abstract

Methods and reagents for diagnosing neoplasia or a susceptibility to neoplasia are provided, which methods comprise detecting the aberrant expression of a gene at the heat shock cognate 73 (HSC73) locus or detecting a mutation at the HSC73 locus. Such methods are useful for example in the detection and/or prediction of carcinomas of the breast.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

CANCER GENE

This invention relates to methods of diagnosing neoplasia or a susceptibility to neoplasia, in particular malignant neoplasia, which method is based on detecting abnormalities in or affecting the HSC73 (Heat Shock Cognate 73) gene (and/or gene products). The invention further relates to reagents for use in such methods. The invention also relates to genetically modified non-human mammals having a general weakness or a predisposition to cancer and/or other diseases, as a result of a genetic modification to or affecting a HSC73 gene.

Cancer results from the accumulation of somatically acquired genetic changes, sometimes associated with inherited predisposing germline mutation. The initiation of tumour development is considered an irreversible genetic change that occurs in cells either spontaneously or upon exposure to chemical and physical agents. A selective expansion of initiated cells then occurs with a concurrent increase in genetic lesions. In the final stage of tumour development, progression, sufficient mutations in oncogenes and tumour suppresser genes accumulate to irreversibly remove the cells from growth control. Mutations which promote malignancy are expected to result in the dysfunction or loss of proteins which protect DNA from damage, proteins which repair DNA and proteins involved in cell cycle checkpoints which control the normal chromosome complement. Therefore, while there is a selection of mutations which remove the normal cell from growth control, many mutations which have no function in tumour promotion will also accumulate as a result of damage to genome maintenance proteins. Identification of mutational events which cause cancer and the characterisation of their gene products will enable better diagnosis and ultimately treatment for cancer sufferers. Identification of cancer-related genes will also enable the development of model systems for testing toxicity (in particular carcinogenicity) of chemicals and biochemicals for use e.g. in pharmaceutical, pesticide and other

compositions.

The kinetics of age incidence of cancer in mice indicate that six or seven gene hits are required to produce the malignant phenotype. This makes any mouse model of carcinogenesis slow and generally dependent on random genotoxic events. A model of susceptibility is required which has a totally defined background. Mice with a specific gene mutation (e.g. *TRP53*), gene amplification (e.g. *Hras*) or expressing a viral gene (e.g. *HBV-X*) are not good models for carcinogenesis testing since a specific pathway which results in the malignant phenotype is compromised. Only those carcinogens which promote cancer via the specific compromised pathway would be detected using such a model.

A mouse is needed which is compromised in all pathways which can result in the malignant phenotype but which is not initiated or committed in any of them. What is needed is a hemizygous mouse which is compromised generally in homeostasis; compromised but not initiated to neoplasia by any specific pathway. Such a mouse model would detect all agents which can promote cancer, in a shorter period than an unmodified mouse, since less gene hits should be required.

Despite the screening programmes in place in recent years, breast cancer remains a major problem, with a death rate of 15,000 women per annum in the United Kingdom. Better screening techniques are therefore needed. Where suspected breast cancer is detected at an early stage, surgical removal can be an effective treatment. However, the two generally accepted potential precursors to breast cancer, ductal carcinoma *in situ* (DCIS) and lobular carcinoma *in situ* (LCIS), only actually result in malignant (invasive) disease in about 25% of DCIS cases and 20% of LCIS cases. Thus, it would be of significant benefit to be able to identify DCIS and LCIS and other lesions which are actual precursors to invasive breast cancer.

For some time, the long arm of chromosome 11 (11q) has been implicated in malignancy. An anonymous locus at 11q22.3 (D11S29)

was found to be deleted in a significant number of breast cancers. Subsequently the incidence of 11q allelic deletions in breast cancer was examined by Southern blot analysis in paired fresh normal and breast cancer tissue (1,2). Of 41 patients with sporadic breast cancer, 59% were shown to have lost one allele at the D11S29 locus. Four other probes (D11Z1, INT2, DRD2, ETS1) proximal and distal to D11S29 did not show allelic deletion. This study was expanded by a microsatellite PCR study. Of 83 patients who were heterozygous at D11S29, 42 % demonstrated allelic imbalance (3).

To finely map the deletion event(s) in breast cancer, paraffin embedded archival tissue was examined by PCR with multiple microsatellite markers. It was found that the region 11q22-q23.1 (8cM) was deleted in 63% of sporadic breast cancers and a novel region of allelic deletion at chromosome 11q25 (2cM) was also identified in 51% of the cases studied (4). In all, deletions along 11q were present in 72% of sporadic breast cancers. When contrasted with a 10% "background" rate of allelic deletion in colorectal carcinomas, these data are strongly indicative of genetic events occurring at chromosome 11q22-q23.1 and 11q25 which may be of importance in the pathogenesis of breast cancer. The deletion event at 11q22-23.1 has been examined in DCIS (a putative pre-invasive lesion) in patients with concomitant invasive cancer. Of 18 patients, 17 were concordant for deletion/or its absence at D11S29 (3). Further allelic deletion at this locus was present in lymph node metastasis (75%) from carcinomas with the deletion. Allelic deletions in invasive breast cancer have been confirmed by others at 11q22-24 (5) or 11q22-23 (6).

Deletion events at 11q22-24 have also been described in 40% of invasive cervical cancers (7) and 67% of matched pairs of normal and melanoma tissue from non familial patients (8). We have also shown the deletion at 11q22-24 in 8 of 108 informative cases of pure CIN3 (cervical intraepithelial neoplasia, Grade III) (9). Given that it takes 5-15

years for CIN3 to progress to invasive cancer (when it does indeed progress), these and the DCIS data indicate that allelic deletion at 11q22-24 may be an early event in the neoplastic process. Deletions of 11q have also been described in ovarian cancer (11q22-24) (10), gastric cancer (11) and lung cancer (12). Collating the allelic deletion data along 11q for all tumour types (see Figure 1), 3 minimal regions of allelic loss have been identified at 11q22-23.1, 11q23.3-24 and 11q25-qter (13). This suggests genetic rearrangement events along chromosome 11q may accumulate affecting up to 3 specific loci during the development of these carcinomas.

Several other lines of evidence implicate chromosome 11q in malignancy. Cytogenetic analysis of breast cancer has demonstrated anomalies at 11q22-25 (14). Numerical aberrations in chromosome 11 and 17, and chromosome 11 and 17 inequality, have been correlated with lymph node metastases or disseminated disease in breast cancer patients (15). Since the 1980's, micro-cell transfer has implicated chromosome 11 as a potent tumour suppresser source in cervical derived HeLa cervical cancer cells (16). Functional tumour suppression with chromosome 11q has also been shown in breast (6), lung (17) and melanoma (18) cell lines and metastasis suppression has been identified in the MDA breast cancer cell line (19).

Thus, although mutations in certain regions of chromosome 11q have been linked to cancer, no basis has been provided for a reliable test relating to this. The regions of the chromosome involved have been until now too large and non-specific. For example, the region of the chromosome from 11q23.3 to 11q25 is greater than 10 megabases long and possibly contains 50 or more genes (most of which are as yet unidentified).

The first murine heat shock cognate sequence published was that of a clone obtained from a teratocarcinoma derived cell line (F9) from the 129/SvJ strain of mouse (27). This sequence was named HSC72. Subsequently a second heat shock cognate sequence was published and

named HSC73 (28). These two sequences differ by a point mutation in exon 3 which results in a single amino change at position 129 from Asp to Asn.

It has now been discovered by the inventors that mutant products of a gene which was originally cloned in 1987 (20), the human Heat Shock Cognate 73 (HSC73), are involved in cancer. HSC73 had previously been located to chromosome 11q23.3-q25 (21) and has in the past been referred to in humans as HSC70 and HSC71. HSC73 is normally present in 2 copies, one on each chromosome 11.

The following observations have been made:

- HSC73 has been fine-mapped to a YAC (yeast artificial chromosome) at 11q23.3.
- HSC73 expression is qualitatively aberrant by Northern blotting, RT-PCR (reverse transcriptase-PCR) and sequencing in 70% of RNA samples extracted from immortalised malignant breast cell lines and non-familial (i.e. sporadic) breast cancers, but not in paired constitutional samples (both breast and skin). Sequencing demonstrates that the aberrant transcripts are generally interstitial deletions of a large number of exons encoding the HSC73 protein (typically covering exon 2 or 3 to exon 9) with break points in exons.
- HSC73 is therefore abberant at the RNA level in 70% of sporadic breast cancers.
- The expression of this gene in breast tissue has also been examined directly by RNA *in situ* hybridisation and immunohistochemistry using polyclonal antibodies raised against synthetic peptides. The tumour cells show either absent or reduced expression of HSC73 RNA in 70% of sporadic breast carcinomas and 50% of DCIS.
- Breast tumours show increased expression of HSC73 protein in approximately 70% of cases (it is believed that this results from the accumulation of mutated protein).
- It is anticipated that the mutated HSC73 RNA is a result of mutated

DNA in these cases and results from point mutations. Evidence for this is discussed herein.

- Disruption of HSC73 gene products has been found in cancers other than breast cancer, for example cervical and colorectal carcinoma and a malignant lymphocytic cell line (for which mutated transcripts have been detected).

Further to these findings we have also discovered DNA mutations in the human HSC73 gene in clinical carcinomas of the breast. This has been achieved by PCR amplification of individual exons of HSC73 at 11q23.3 using oligonucleotide sequences within the introns of the gene from DNA extracted from clinical carcinomas of the breast. Pseudogenes have not been amplified since the oligonucleotide sequences are specific for the genomic HSC73 locus at 11q23.3. All amplified DNA fragments have been sequenced. Intron sequences confirmed the fragments were genomic sequences of HSC73 at 11q23.3.

The DNA mutations so far identified in HSC73 in clinical carcinomas of the breast are generally point mutations and small insertions/deletions of sequence. The mutations are described in Example 1. The mutations identified in HSC73 are in exon 3 (the second coding exon) of the gene and are not present in DNA purified from adjacent non-malignant breast tissue or skin from the same patient. These nucleotide changes result in amino acid substitutions. The nucleotide changes and amino acid substitutions are listed in Example 1. This suggests that some change(s) in HSC73 protein function (either gain or loss) is acquired as a result of mutation in exon 3 which confers a growth advantage to the tumour cell. Exon 3 is critical for the ATPase activity of the protein (O'Brien *et al* 1996, J. Biol. Chem. 271; 15874-8).

Allelic imbalance data implicates 11q23.3 as the location of a multi-tumour suppressor gene. The HSC73 gene maps to this location and is mutated in human clinical carcinomas of the breast.

Disruption of HSC73 is expected to be involved in a wide

variety of malignant and non-malignant diseases. The protein encoded by HSC73 is an essential component of the centrosome (22), a genome caretaker (23), a participant in protein degradation pathways (24 – subsequently independently confirmed) and essential for maintaining proteins in a semi folded state in order to enable translocation through the mitochondrial and endoplasmic reticulum (it is therefore essential for steroid hormone receptor activation) (25).

The deletions in the RNA encoding HSC73 which have been characterised in malignant disease also delete three small nucleolar RNAs 14 (SNURPs) which are essential for ribosome biosynthesis and 18S rRNA processing (26). The three SNURPs are encoded by short sequences in introns 5, 6 and 8 of HSC73.

The invention provides in one aspect a method of diagnosis performed on a biological sample, which method comprises detecting the aberrant expression of a gene at the HSC73 locus, or detecting a mutation at the HSC73 locus, in the sample.

Such a method is useful in particular for diagnosing neoplasia or a susceptibility to neoplasia, especially malignant neoplasia. The diagnosis may be a predictive diagnosis, for example it may predict whether or not *in situ* neoplasia such as DCIS will become invasive.

In further aspects, the invention provides novel reagents for use in the method according to the invention. Such reagents include but are not limited to:

i) Oligonucleotides and panels of oligonucleotides capable of indicating the presence of a mutation at the HSC73 locus by hybridising or by not hybridising to DNA or RNA. Such oligonucleotides may be employed as primers or probes. The invention is concerned in particular with oligonucleotides which hybridise specifically under stringent conditions, to exon 3 or a region of exon 3 of the HSC73 gene. The oligonucleotides may be specifically complementary either to the wild type HSC73 sequence or to a mutant HSC73 sequence.

ii) Monoclonal antibodies and panels of monoclonal antibodies capable of indicating the presence of a protein product of a mutant HSC73. The invention is concerned in particular with antibodies which bind specifically to a region of the HSC73 protein encoded by exon 3, especially mutated exon 3. Such antibodies can be easily raised for example using recombinant HSC73 polypeptides or peptides generated using known techniques.

The invention further provides synthetic peptides capable of generating such antibodies and isolated or recombinant proteins or polypeptides capable of generating such antibodies.

In still another aspect, the invention provides a genetically modified non-human mammal having genetic weakness to disease, in which the somatic cells and germ cells have at least one mutated or deleted copy of HSC73. Such genetically modified animals will be susceptible to cancer by a variety of pathways, since their control of homeostasis is weak. At the neonatal stage, they will generally be healthy animals which are not initiated in any cancer pathway.

The invention further provides an embryonic stem cell transformed with a nucleic acid capable of replacing a copy of HSC73 in the cell. The nucleic acid may be either a mutated copy of HSC73, or a marker such as a drug resistance marker. Such a cell may be regenerated into a genetically modified animal according to the invention.

The invention provides in yet another aspect a cell line genetically engineered to have at least one mutated or deleted copy of HSC73. Genetically modified HSC73 cell lines will be useful as targets for therapeutic testing and development of anti-cancer agents.

In another aspect the invention provides a method of testing a compound for biological activity, particularly toxicity, which method comprises exposing a cell line or a genetically modified non-human mammal according to the invention, to the compound.

The invention also provides kits for diagnostic and screening

purposes, in particular a kit for performing a method of diagnosis described herein which kit comprises an oligonucleotide which hybridises specifically to a DNA which encodes a mutated or unusual polymorphic form of an HSC73 mRNA or peptide, and not to a DNA which encodes wild type HSC73 mRNA or peptide. Alternative kits are provided for detecting aberrant HSC73 protein and mRNA levels as described herein.

Furthermore, the invention provides uses in gene therapy. Thus, the invention provides a method of treatment comprising introducing into a patient at least one copy of a nucleic acid which encodes a native HSC73 under the control of a suitable promoter; and a composition comprising a gene delivery vector which comprises a nucleic acid encoding an HSC73 protein under the control of a suitable promoter, together with a pharmaceutically acceptable carrier. Suitable vectors for delivery of the nucleic acid are well known in the art of gene therapy, as also are appropriate promoters for the expression of the HSC73 protein *in vivo*. The introduction of a gene in gene therapy may be carried out on cells *in vitro* which cells are then introduced or reintroduced into the patient, or it may be carried out by delivery of the gene to cells *in vivo*.

In addition, the invention will be useful in the development of new anti-cancer therapies. Pharmaceutical agents which can upregulate or replace a function of HSC73 or which are inhibitory to mutated HSC73 may be beneficial. Mutated HSC73 may indicate tumour susceptibility to individualised therapeutic protocols. Gene or protein replacement therapy for HSC73 in malignant tissue may also be beneficial, since HSC73 is ubiquitously expressed. A suitable experimental model for the technical development of gene therapy would be a genetically modified mouse or other non-human mammal model, such as a hemizygous HSC73 "knockout" mouse model. As well as novel therapies directed to DNA (gene therapy), therapies directed to RNA (antisense technology) and proteins (binding, modifying agents and immunotherapy) are also feasible.

The human HSC73 sequence is found at GenBank Accession

Number Y00371 (se Figure 5). The murine HSC73 sequence is found at GenBank Accession Number J17982.

Figure 1 shows an ideogram of chromosome 11, focusing on the 11q22-qter region. Marker loci are arrayed vertically (inter marker distance not to scale) and tumour types listed horizontally. Thick vertical lines indicate minimum regions of deletion reported (references in text), and barred boxes show consensus regions of deletion between different reports, for a given tumour type. Regions 1 and 2 show possible localisation of multi-tumour suppressor genes deleted in above tumour types. HSC73 maps to region 2.

The invention will now be described in more detail below. It will be understood that the following description is illustrative and is not intended to limit the scope of the invention in any way.

Diagnostic and Prognostic Techniques

Some of the possible techniques and reagents for use in the invention are as follows:

1) Immunohistochemistry (on fresh and formalin fixed paraffin embedded tissue)

The gene which encodes HSC73 has nine exons, eight of which are coding. The gene is mutated at the DNA and/or the RNA level in a high percentage of cases of several types of tumour (see below). This leads to aberrant expression of HSC73. We have characterised the expression of HSC73 protein in malignant disease of the breast, ovary, prostate and lung using a commercially available goat polyclonal antibody raised against a synthetic peptide which bridges exons eight and nine of the HSC73 gene (Santa Cruz Biotechnology, Inc. product HSC70 (K19): catalogue number sc-1059). The expression of HSC73 in the tumours listed detected with the polyclonal antibody, is increased over normal tissue in approximately 80% of cases. The expression of HSC73 in other malignant disease (e.g. cervix, colon, stomach and skin) can be easily

characterised using this antibody and correlated with diagnostic and prognostic data.

Further, since the gene encoding HSC73 can be mutated both at the DNA and/or RNA level, a panel of monoclonal antibodies raised against synthetic peptides with one antibody against each coding exon, can be used to characterise the expression of HSC73 in malignant tissue. The expression of HSC73 detected with each of these antibodies is correlated with diagnostic and prognostic data and with the expression detected by other antibodies. Increased expression of the protein detected with an antibody against one exon may reflect an increased expression of the gene from one (or both) alleles or an accumulation of mutated HSC73 which cannot be degraded correctly. A reduction in expression detected by an antibody may reflect a deletion in the gene either at the DNA and/or RNA level which removes an amino acid sequence from HSC73 expressed from one or both alleles; or it may reflect deletion of the entire gene from one (or possibly both) alleles.

The full sequences for both copies of the gene in malignant breast tissue and in paired normal tissue can be obtained by standard techniques. If a conserved point mutation(s) is detected which affects protein sequence(s), a synthetic peptide(s) corresponding to this sequence may be prepared and a monoclonal antibody raised against it which is specific to the mutant protein. The expression of HSC73 detected with one or more such antibodies is then correlated with diagnostic and prognostic data.

It is anticipated that an antibody or a combination of such antibodies may be used for detection of aberrant HSC73 expression which will be of use in a routine histopathology laboratory for:

- Diagnosis of neoplasia (invasive and *in situ*)
- Prognosis of malignancy (invasive disease).
- Prognosis of *in situ* neoplasias such as DCIS/LCIS and CIN3.

2) RNA *in situ* hybridisation (on fresh and formalin fixed paraffin embedded tissue).

We have characterised the expression of HSC73 mRNA in malignant disease of the breast, ovary, prostate and lung using a synthetic antisense oligonucleotide complementary to exon nine. The expression of HSC73 mRNA in these tumours detected with this synthetic oligonucleotide is decreased in comparison to normal tissue in approximately 80% of cases. The expression of HSC73 mRNA in other malignant disease (e.g. cervix, colon, stomach and skin) can be easily characterised using this synthetic oligonucleotide and correlated with diagnostic and prognostic data.

Further, since the gene encoding HSC73 can be mutated both at the DNA and/or RNA level, an antisense oligonucleotide against each coding exon can be used to characterise the expression of HSC73 mRNA in malignant tissue from all tissues. The expression of HSC73 mRNA detected with each of these synthetic oligonucleotides is correlated with diagnostic and prognostic data and with the other oligonucleotide. Increased expression of the RNA detected with an oligonucleotide against one exon may reflect an increased expression of the gene from one (or both) alleles or it may reflect an accumulation of mutated HSC73 mRNA which cannot be degraded correctly. A reduction in expression detected by a synthetic oligonucleotide may reflect mutation in the gene at the DNA and/or RNA level, which removes the target sequence from the HSC73 mRNA expressed from one or both alleles. Alternatively the mRNA detected may be decreased because it has a decreased half life as the result of a mutation or because it is not being expressed at its usual levels due to errors in the control of its transcription which may involve a positive feedback loop. The gene may also be entirely deleted from one (or possibly both) alleles.

If, in sequencing the HSC73 gene in malignant breast tissue and paired normal tissue, a conserved point mutation(s) is detected which

affects mRNA sequence(s), a synthetic oligonucleotide(s) specific to it can be prepared. The expression of HSC73 mRNA detected with each of these synthetic oligonucleotides is then correlated with diagnostic and prognostic data.

Synthetic oligonucleotides may be prepared which are antisense to each of the three U14 SNURPS, for use in examining the expression of the three SNURPS in malignant tissue from all tissues. The expression of the three SNURPS detected using these synthetic oligonucleotides is then correlated with diagnostic and prognostic data.

A synthetic oligonucleotide staining procedure for HSC73 mRNA expression will be of use in a routine histopathology laboratory for:

- Diagnosis of neoplasia (invasive and *in situ*)
- Prognosis of malignancy.
- Prognosis of *in situ* neoplasias such as DCIS/LCIS and CIN3.

The pattern of expression of HSC73 protein and HSC73 mRNA in malignant disease largely parallels that of p53 and TP53 mRNA (29). Staining techniques similar to those for p53 and TP53 mRNA can be devised for HSC73 and HSC73 mRNA.

3) <u>DNA analysis</u>

Sequence analysis of the HSC73 gene has revealed mutations as described above in malignant breast tissue and paired normal tissue shows no corresponding mutation.

A suitable method for detecting the numbers of copies of each HSC73 exon and each of the three SNURPS is as follows. At least two oligonucleotides, each labelled with a different detectable label, are provided. One oligonucleotide hybridises to a reference marker in the genome which is known not to be mutated in malignant disease. The second oligonucleotide hybridises to an exon of HSC73 or the three SNURPS. Comparison of the signals detected allows gene dosage measurements to be made for each exon of HSC73 and each of the

SNURPS. Gene dosage measurements are correlated with diagnostic and prognostic data. The detection systems that may be used include radiolabelled or fluorescence labelled oligonucleotide probes (including prompt [e.g. fluorescein] and time resolved fluorescence [lanthanide chelates like Europium and Samarium]) with or without a prior gene amplification (e.g. by PCR) step.

Detection of point mutations may also involve oligonucleotide hybridisation. The oligonucleotides used for these experiments will normally be very short, and will be specific for the mutation. They will be designed such that they will not hybridise to the published normal sequence of the gene under stringent conditions. Point mutations are correlated with diagnostic and prognostic data. Detection systems may be as already described.

Suitable methods to detect mutation also include technologies such as PCR amplification, cycle sequencing and Taqman™ PCR. Mutations will again be correlated with diagnostic and prognostic data.

Detection of mutations (such as deletions or point mutations) individually and in combination can be used for:

- Diagnosis of neoplasia (invasive and *in situ*)
- Prognosis of malignancy (invasive disease).
- Prognosis of *in situ* neoplasias.

Mutations within the HSC73 coding region of the genome may be germline and therefore transmitted through the population predisposing individuals to cancer (and possibly other disease).

4) <u>RNA analysis</u>

Sequence analysis of the HSC73 mRNA in malignant breast tissue and paired normal tissue shows mutation. To date the majority of mutations in the HSC73 mRNA characterised are deletions in the HSC73 mRNA. Splice variation is also a possibility. All mutations detected will be correlated with diagnostic and prognostic data.

A suitable method for detecting the numbers of copies of each HSC73 exon expressed in the mRNA and each of the three SNURPS is similar to that described above under "DNA analysis". At least two oligonucleotides labelled with different detectable labels are provided. Comparison of the signals detected allows RNA dosage measurements to be made for each exon of HSC73 and each of the SNURPS. mRNA dosage measurements are correlated with diagnostic and prognostic data. Detection systems may be as already described for the DNA analysis techniques.

Point mutations may be detected using short oligonucleotide probes specific for the mutation. The probes are designed such that they will not hybridise to the published normal sequence of the gene under stringent conditions. Point mutations will be correlated with diagnostic and prognostic data. Detection systems used may be as already described.

Suitable methods to detect mutations also include technologies such as PCR amplification, cycle sequencing and Taqman™ PCR. Mutations will again be correlated with diagnostic and prognostic data.

Detection of mutations in the mRNA or SNURPS (such as deletions or point mutations) individually and in combination may be used for:

- Diagnosis of neoplasia (invasive and *in situ*)
- Prognosis of malignancy (invasive disease).
- Prognosis of *in situ* neoplasias.

<u>**Model Systems for Toxicity Testing**</u>

A mouse carrying a hemizygous deletion of HSC73 and U14 SNURPS is a good model for carcinogenesis testing since:

- General house keeping is compromised.
- The HSC73 protein is expressed in every cell in the body (murine and human).

- HSC73/U14 SNURPS are mutated both at the RNA and protein level in many different human cancers.

A mouse carrying a hemizygous deletion of HSC73 and U14 SNURPS is susceptible to cancer in many ways:

- Increased mutation rate since there is reduced genome protection as a result of a reduced chaperone complement.
- Compromised centrosomes and cytokinesis compromise mitosis in general. HSC73 is an essential component of the centrosome.
- Steroid hormone activation is compromised leading to aberrant inter and intra cell signalling and possible increased mitogenesis.
- Protein degradation in proteosomes is compromised since HSC73 is essential for the transport of a subset of proteins into the 28S proteosome.
- The correct functioning of all cell surface molecules is compromised since HSC73 is essential for transport across membranes. Cells are therefore compromised in regard to self/self and self/antigen recognition.
- Ribosome biosynthesis and RNA processing are compromised since U14 SNURPS are essential for pre-ribosomal RNA splicing. As a result translation is generally compromised.

A mouse carrying a hemizygous deletion of HSC73 and U14 SNURPS is a good model for carcinogenesis testing. The pathway to neoplastic disease promoted by a chemical in this animal may be detectable. As the hemizygous deletion of HSC73 and U14 SNURPS knock out mouse is characterised with respect to development the compromised cellular pathways and functions will be distinguished. It should be possible to demonstrate that a chemical is a carcinogen in the hemizygous deletion of HSC73 and U14 SNURPS mouse model of carcinogenesis testing if for example it damaged the remaining steroid hormone activation function to an extent which removed hepatocytes from normal growth control.

Manufacture of a mouse carrying a hemizygous deletion of HSC73 and U14 SNURPS.

The cloning of murine HSC73 has been described previously (27,28).

Isolation of a 129/SvJ genomic BAC HSC73 clone requires PCR screening of the library with oligonucleotides which will not hybridise to processed pseudogenes elsewhere in the genome. The best oligonucleotides therefore span the U14 SNURPS which are in the introns of HSC73 in mouse and man. This provides three autonomous reactions with which to screen the library.

Isogenic targeting constructs are then prepared from the BAC clone for homologous recombination. In these clones the entire HSC73 gene and associated SNURPS may be replaced by neomycin or another drug marker or the "normal" HSC73 may be replaced by a mutated copy and the drug marker inserted into the BAC elsewhere.

The homologous recombination HSC73 clone is electroporated into RW-4 embryonic stem cells and plated onto MEF feeder cells. Transformed embryonic stem cells are selected for by antibiotic selection, and a minimum of 150 clones are screened by Southern blotting and PCR for appropriate homologous recombination between the engineered target and the "normal" HSC73 locus in the 129/SvJ mouse genome. Recombinant clones are expanded and karyotyped. Those in which satisfactory recombination has occurred are used for either blastocyst injection or morula aggregation in order to generate chimeric mice. These will be used to obtain founder mice. The mice strains will be maintained by conventional breeding.

EXAMPLES

EXAMPLE 1

HSC73 expression in breast cancer was investigated by independent RNA in-situ hybridisation (RNA-ISH) and protein immunohistochemical (IHC) analysis of 8 non-malignant (3 fibrocystic change, 5 fibroadenomas) and 44 invasive breast cancer cases.

On in-situ analysis, HSC73 RNA was expressed at a high level (RNA-ISH score 5/5) in all normal tissues examined (breast acini, lymphocytes, endothelium, skin). This was associated with moderate tissue levels of the HSC73 protein (IHC score 3/5). Assay of the 8 non malignant breast lesions revealed a similar uniform intense staining of HSC73 RNA in the lesional cells and adjacent histologically normal breast acini (RNA-ISH score 5/5). Increased protein staining (IHC score 4/5) was seen in 3 of 5 fibroadenomas (60%) and focally in the 3 fibrocystic samples (apocrine metaplastic cells only) (Figure 1.1A and B). As expected of a constitutive 'housekeeping' gene, the ~2.3kb HSC73 mRNA transcript is uniformly expressed in many normal tissues (Figure 3 a).

Of 44 archival breast cancers (and adjoining normal acini) analysed, 32 cases (73%) showed a moderate/marked reduction in HSC73 mRNA expression (RNA-ISH score 1-3/5), and 31 of these (70% of total) had increased expression of the HSC73 protein (IHC score 4-5/5) (Figure 2.3 A and B). Tumour infiltrating macrophages and lymphocytes had high levels of both HSC73 RNA and protein. Metastatic tumours from the same cases showed the inverse relationship between mRNA and protein strikingly, with near total absence of the HSC73 RNA and elevated HSC73 protein (Figure 2.4 A and B). In a series of over 40 cases of pure ductal carcinoma in situ (DCIS), ~50% have this inverse microscopic pattern of RNA and protein expression (Figure 2.2 A and B) (NCAH *et al*, m/s in preparation). This possibly indicated a biologically important difference between malignant and non-malignant cells in the HSC73 gene products.

18 cases (and adjacent normal tissue) from the above series,

with material for RNA and DNA analysis, were subsequently used to independently assess gene structure (9 cases also had sufficient tissue for Western blotting). Aberrant bands were visualised in 8 of 18 tumours (44%) on northern blots (the gene has no splice variants described), indicating that they comprised a significant fraction of the HSC73 RNA (Figure 3 b). Of the 18, 12 cases (67%) had intra-exon breakpoints (sites at which a deletion in the mRNA has occurred) in the HSC73 gene by RT-PCR sequencing (in 3 normal lanes, aberrant RT-PCR bands were present; 2 were aberrantly spliced at exon-intron boundaries with exons deleted, and one with an intra-exon breakpoint actually had residual tumour) (Figure 3 c). Further, Western blots showed aberrant truncated protein in 7 of the 9 cases, evidence that the mutant RNA is translated to abnormal protein (Figure 3 e). Given the highly conserved nature of the HSP70 gene family (especially at the N-terminal end of HSC73), the probes used for RNA, protein and DNA analysis could not span the complete sequence, necessarily limiting mutation detection and data overlap from different analytic methods. Nonetheless, all cases with truncated bands on Northerns had RT-PCR mutations, and where no aberrant band was detected by Northern analysis and RT-PCR, no aberrant bands were detected by Western blotting.

Analysis of 8 breast cancer (and one mononuclear) cell lines was performed at RNA, DNA and protein level. Northern analysis revealed aberrant transcripts in MDA-MB435, MDA-MB468 and the U937 ('macrophage') cell lines (Table below). No mutations were detected in MDA-MB157, MDA-MB231, MDA-MB361, MDA-MB453, SKBR3 or T47D (data not shown). Mutations in MDA-MB435, MDA-MB468 and U937 were seen by RT-PCR (Figure 3 d) and sequenced. Western blot analysis demonstrated a truncated HSC73 isoform in the U937 line (Figure 3 f).

To summarise, aberrant inverted patterns of HSC73 mRNA and protein expression occurred in 31 of 44 (70%) invasive breast cancers. Aberrant HSC73 bands were also detected on Northern blots (44%) and

RT-PCR (67%). Sequencing revealed intra-exon breakpoints in most of the truncated tumour RNA, while Western blots and IHC showed aberrant protein expression patterns and mutation, with over-expression of wild-type (Wt) band and the presence of multiple truncated mutant (Mt) bands. Mutations in the DNA have been discovered as described above.

Results are displayed in the following table and in Figures 2 and 3.

TABLE

Case	Northern	RT-PCR breakpoint(s)	Location of breakpoints	RNA-ISH (N score = 5)	IHC (N score = 3)
16	Mt	1114-4712	Exon 2 - 3' untranslated	4	4-5
70	Mt	1046-4714	Exon 3 - 3' untranslated	4	4-5
22	Mt	1593-4683	Exon 3 - 3' untranslated	3	5
24	Mt	1117-4763	Exon 2 - 3' untranslated	2-3	4-5
34	Mt	1158-4751	Exon 2 - 3' untranslated	2-3	4-5
38	Wt	1595-4716	Exon 3 - 3' untranslated	3	5
54	Mt	1085-4688, 1639-4608	Exon 2 - 3' untranslated, Exon 3 – Exon 9	2	5
58	Wt	1166-4725, 2826-4628	Exon 2 - 3' untranslated Exon 5 - 3' untranslated	1-2	4-5
20	Mt	1639-4608	Exon 3 – Exon 9	3	4
52	Wt	1163-4751	Exon 2 - 3' untranslated	3	4-5
62*	Wt	1648-4688	Exon 3 - 3' untranslated	3	4-5
64	Mt	1038-4715	Exon 2 - 3' untranslated	3-4	5
66	Wt	-	-	3-4	4-5
60	Wt	-	-	3-4	4-5
26	Wt	-	-	5	3
36	Wt	-	-	5	3
50	Wt	-	-	5	3
72	Wt	-	-	5	3
MDA-MB435	Mt	1191-4753	Exon 2 - 3' untranslated	-	-
MDA-MB468	Mt	Splice variant	Exon 2 – Exon 9	-	-
U937	Mt	2185-4648	Exon 4 - 3' untranslated	-	-

The table shows deduced HSC73 gene structure and expression data for the 18 fresh cases, with mutated cell lines listed below. N= normal; Nd=not determined; Mt=mutation; Wt=wild type. * tumour with microscopic invasion into adjacent 'normal'. Note that all aberrant bands were not completely sequenced once a breakpoint was determined. Larger deletions are thus over-represented.

Methods:

HSC73 mapping

A contiguous chromosome 11q22-qterm YAC library plated in an ordered array was hybridised to oligonucleotides specific for HSC73, localised to an 11q23.3 YAC (y667F08) between D11S1336 and D11S1284 (11.7cR; ~500kb) and confirmed by PCR and southern analysis (data not shown).

Cases

44 archival paraffin embedded blocks of sporadic breast cancer and paired normal tissue were serially sectioned and analysed. Of these, 18 cases had frozen tissue for RNA, DNA and protein (9 cases) extraction.

RNA analysis:

RNA-ISH: sections were dewaxed, rehydrated, proteinase K digested (1mg/ml in 2mM $CaCl_2$, 10mM Tris, pH 7.4) for 30min at 37°C and hybridised overnight at 37°C in 50% deionised formamide, 2X SSC, 20% dextran sulphate, 100mM antisense probe (5' CATTCCTCCTGGCATGCCTCCTGCACT-3' [SEQ ID NO: 1]), digoxigenin labelled (Boehringer). Sections were washed in TBS, pH 7.6, 0.1% Triton X-100, blocked (TBS, pH 7.6, 0.1% Triton X-100, 3% BSA, 20% goat serum), incubated with 1: 500 alkaline phophatase conjugated anti-digoxigenin antibody (Boehringer) for 30 mins and washed again in TBS (pH 7.6, then pH 9.0) for 5 mins. Signal was developed in 100mM Tris, pH 9.0, 50mM $MgCl_2$, 100mM NaCl, 330mg/ml nitroblue tetrazolium (NBT), 165mg/ml 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and 1mM levamisole for up to 30 mins. Control sections included a complementary sense probe, RNase pre-treatment and competitive signal attenuation by unlabelled antisense probe.

Northerns: 20mg total RNA was resolved on 1.5% agarose-3% formaldehyde gels, integrity visualised by ethidium bromide, blotted to Hybond N+ (Amersham) and hybridised overnight to a $\alpha^{32}P$ labelled cDNA probe from exon 5-3' polydenylation site (ATCC IB743), in 50% formamide, 0.75M NaCl, 36mM Na_2HPO_4, 4mM NaH_2PO_4, 0.15M Tris-HCl, pH 8.0, 10mM EDTA, 5X Denhardts, 200µg/ml denatured salmon sperm DNA and 0.5% SDS (w/v) at 42°C. Washes were for 15 min in 2X SSC, 0.1% SDS; 0.5X SSC, 0.1% SDS and up to 4 times in 0.1X SSC, 0.1% SDS at 65°C and autoradiographed on Xomat-MS film (Kodak) at -70°C for up to 7 days. Multiple-tissue northerns were purchased (Clontech) and probed as above.

RT-PCR and cDNA sequencing: 10µg total RNA was incubated with 10units RNase-free DNase I (Boehringer) at 37°C for 45 min, twice phenol-chloroform extracted and ethanol precipitated. cDNA was transcribed by SuperScript II RNase H^- reverse transcriptase (Life Technologies) with the HSC73 specific primer (5' AAGACATTGCATTTTCCACTTACAATA 3' [SEQ ID NO:2]). 1.5µl cDNA was used in 35 cycles of hot-start PCR using 10pmoles of primers P1 (5' TTGTGGCTTCCTTCGTTA 3' [SEQ ID NO:3]) and P2 (5' CAGTGTATAAAGTGCAATG 3' [SEQ ID NO:4]) in 200µM dNTPs, 1.5mM $Mg(OAc)_2$, 1X XL PCR buffer (Perkin-Elmer) and 0.5 units of rTth XL Taq (Perkin-Elmer) in a 25µl reaction on a Perkin Elmer 9600 cycler at 94°C denaturation for 0.25 min, 54°C annealing for 0.25 min and 72°C extension for 1.25-2.75 min. 1µl of 1:50 PCR products were used in 35 cycles of another hot-start PCR with 10pmoles of the nested primers P3 (5': TTGGAGCCAGGCCTACAC 3' [SEQ ID NO:5]) and P4 (5' CCTTCCCCTGTGCATATG 3' [SEQ ID NO: 6]) in 200µM dNTPs, 3mM $MgCl_2$, 50mM KCl, 10mM Tris-HCl (pH 7.5), 0.1% Triton X-100 and 0.5 units SuperTaq (HT Biotechnology, UK) in a 25µl reaction, cycled similarly with annealing and extension at 65°C and 70°C respectively. PCR products were visualised on 1% agarose gels. Control RT negative samples were blank. DNA fragments were excised, purified with a QIA

quick kit (Qiagen), $\alpha^{35}S$ cycle sequenced by the Amplicycle kit (Perkin Elmer) with each inner primer for 25-30 cycles, resolved on 6% denaturing urea-polyacrylamide gels and visualised on Xomat-AR film (Kodak).

DNA analysis

Southerns: 20µg DNA incubated with *Bam*H1 and *Bgl*II (MBI Fermentas) was resolved on 0.8% agarose gels, filters hybridised to $\alpha^{32}P$ labelled probes subcloned into pGEM (5' probe: PCR generated 700bp intron 1 fragment; 3'probe: 700bp exon 8-3'untranslated) in 50% formamide buffer at 42°C o/n, washed at 65°C in 0.1X SSC, 1% SDS twice for 20 minutes and visualised on Xomat-AR film (Kodak) overnight at -70°C. The 5' probe yielded a single 8.4 kb band, consistent with a single copy of HSC73 in the genome. The 3' probe yielded 7 bands, identifying HSC73 and family members.

PCR: 50ng RNaseA treated DNA was used to amplify ~3.8kb of the HSC73 genomic sequence (intron 1-3' untranslated) by 42 cycles of hot-start PCR with 15pmoles of the primers P5 (5': TGGTTAAGTGTTCTGTTAAG 3' [SEQ ID NO: 7]) and P6 (5' CAATTGTATGGTGCCAATTT 3' [SEQ ID NO: 8]) in 200µM dNTPs, 1.5mM $MgCl_2$, 50mM KCl, 10mM Tris-HCl (pH 8.0), 0.1% Triton X-100 and 0.5units SuperTaq (HT Biotechnology, UK) in a 25µl reaction at 94°C denaturation for 0.25 min, 56°C annealing for 0.25 min and 70°C extension for 2.5-4 min. PCR products were visualised on 0.7% agarose gels.

Protein analysis

Immunohistochemistry: Sections were dewaxed, rehydrated, pressure cooked for antigen retrieval (1.5 mins at maximum pressure; 10mM sodium citrate buffer, pH 6·0), incubated with a goat polyclonal antibody (Santa Cruz, USA) against HSC 73 amino acids 583-601, at 1:200 dilution in PBS/BSA for 60 minutes at 37°C. Secondary incubation used a

biotinylated anti-goat antibody (DAKO) followed by a standard avidin-biotin complex (ABC, DAKO) for 30 minutes and developed with DAB.

Western *Blots*: 10μm cryosections were boiled in 24 mM Tris-HCl, pH 6.8, 4% SDS (w/v), 10% β-mercaptoethanol, 20% glycerol for 10 mins, sonicated, debris pelleted and protein estimated (Bio-Rad). 10μg protein was resolved on 10% polyacrylamide gels, electroblotted onto Hybond C+ (Amersham), blocked (5% (w/v) milk powder in 1X TBS, 0.1% Tween 20), treated with the HSC73 antibody (1:2000) and washed in 1X TBS, 0.1% Tween 20 (v/v). The membrane was incubated in horse radish peroxidase conjugated anti-goat antibody (1:500) (DAKO), washed in 1X TBS, 0.1% Tween 20 (v/v) and developed by ECL (Amersham). Coomassie Brilliant Blue gel staining and a monoclonal antibody against CD68 (KP1) controlled for protein integrity.

EXAMPLE 2

Sequence analysis of exon 3.

Analyses of cell lines and clinical material have revealed mutations including base pair substitutions or point mutations in exon 3 and a two base pair deletion in intron 2. Of 14 tumours, 3 (21%) had point mutations which changed the coding sequence of exon 3 (adjacent normal breast did not have the same mutation). These three cases were all ones in which aberrant splicing was previously detected. Of the three tumours in which point mutations were identified, two were also homozygous and one heterozygous for a 2bp germline deletion in intron 2 adjacent to exon 3. In total, 15% of breast tumours were heterozygous and 15% homozygous for the intron 2 deletion. The intron 2 deletion has not been identified in tumours in which there was no aberrant splicing.

However, blood sampled from 21 women who developed benign breast disease and 12 women who developed malignant breast disease was heterozygous for the intron 2 polymorphism in 30% of each

group.

The polymorphism in intron 2 of the gene may cause susceptibility to aberrant splicing of *HSC*73 in malignant cells; additional factors must be involved since aberrant transcripts have not been observed in normal breast. RNA truncations could occur if cryptic splice sites are produced or if mutated (and/or polymorphic) pre-mRNA primary structure affects tertiary structure bringing incorrect sequences into contact which may be spliced together. The U14 SNURPs which are transcribed via an intron processing pathway concurrent with *HSC*73 mRNA splicing, and the repetitive nature of the intron/exon structure, may make *HSC*73 pre-mRNA splicing inherently susceptible to incorrect splicing in a malignant cell.

The point mutations we have detected in exon 3 of the human tumours which change coding are:

Valine 83 is changed to alanine
gtc is changed to gcc

Glycine 99 is changed to Serine
ggc is changed to agc

Tyrosine 107 is changed to histidine
tac is changed to cac

The polymorphism in intron 2 is:
aaagtgttttcattgacaccttttacagAT
which becomes
aaagttttcattgacaccttttacagAT
where AT is the start of exon 3
the sequence changes 23bp upstream of exon 3.

EXAMPLE 3

Transfection Studies on HSC73

Functional studies were performed using mda-mb435, a breast cancer cell line known to have aberrant hsc73 mRNA, as a model. The HSC73 gene was transfected into the cell line. Three stable subclones of this cell line expressing normal hsc73 from a genomic clone were produced and investigated for *in vitro* and *in vivo* tumour formation in mouse models. Western blotting confirmed increased full length hsc73 protein in all three subcloned cell lines. The clones were each introduced into three test animals. There were three control animals injected with untransfected cells of a breast cancer cell line. Proliferation assays demonstrated that expression of wild type hsc73 in the mda-mb435 cell lines decreases the rate of cell growth both *in vitro* and *in vivo* in nude athymic mice. *In vitro* proliferation rates for the three stable hsc73 expressing subclones and appropriate control cell lines were measured. The stable transfectants divide at approximately ¼ of the rate of division of the wild type cell line (data not shown). *In vivo* proliferation assays showed that 500,000 transfected cells produced no tumours in nude mice (except in a single case) while the wild type mda-mb435 produced tumours in 100% of cases (Figure 7). It is likely that in the case where tumours did develop from the transfected cells, the HSC73 gene was for some reason no longer being expressed (see figure legend).

In a similar *in vivo* experiment the procedure described was repeated except that 2 x 10^6 cells were injected into each animal and control transfectants expressing the hygromycin resistance gene without HSC73 were included. In all nine mice injected with the stable transfectants (i.e. three mice for each of the three clones) there was no tumour formation. In the hygromycin resistance gene control mice, tumour formation was at the same level as for the mice injected with the wild type cell line.

These results demonstrate that hsc73 can suppress tumour formation. hsc73 expressed in certain breast cells can reduce their

proliferation. There are two likely biochemical explanations for these observations.

1. Mutation in hsc73 prevents apoptosis in breast carcinoma cells. Replacement of hsc73 function in the mda-md435 following stable transfection repairs the apoptotic function and gross proliferation of the cells is therefore reduced. Examples of biochemical defects could include the degradation of c-fos in the 26S proteasome, a known early reaction in the cascade for the activation of caspase 3 by caspase 9 and cytochrome c. Other explanations include the correct cellular localisation of apoptosis related proteins such as BAG and BUZ.

2. mda-mb435 expressing normal hsc73 could be stalled in a stage of their cell cycle progression. The cell line could have impaired wild type hsc73 function as a means of overcoming such a cell cycle block. Obvious classes of molecules to examine for their stability and localisation in the cell lines include cyclins and cdk inhibitors.

Given the pleiotrophism of this gene it may well be that both mechanisms play a role in suppressing malignancy in the wild type protein.

Figure Legends:

Figure 2- Paired serial sections of breast lesions stained for IHC (panel **A**) and RNA-ISH (panel **B**). Normal/non-malignant cells ( ) and malignant cells ( ) are indicated. (**1**) The fibroadenoma shows intense signal in epithelial cells on both analyses. (**2**) DCIS shows increased tumour cell HSC73 protein but low RNA compared to normal acini. (**3**) Invasive breast cancer cells show a similar phenomenon, with compressed normal ducts staining weakly for protein but intensely for RNA, compared to tumour cells with the opposite pattern. (**4**) Metastatic tumour in the lymph node stains highly for protein while RNA signal is virtually absent, contrasting sharply with the surrounding normal lymphocytes. Scale bar = 100μm.

Figure 3- a) Multiple tissue northern blot (l-r: spleen, thymus, prostate,

testis, ovary, small intestine, colon, leucocyte) probed with 1.1kb cDNA probe (with β-actin control panel below) shows uniform expression of HSC73. **b**) Northern of paired normal (**N**) and tumour (**T**) samples (l-r: tumour cases 16, 20, 34, 26, 22) and cell lines MDA-MB435 (1), MDA-MB468 (3) and U937 (5) are shown with aberrant transcripts indicated (◄). **c**) HSC73 RT-PCR samples (l-r: tumour cases 20, 64, 54, 24, 22, 62, 58, 26, 34) with multiple aberrant bands in tumour lanes (**N*** indicates a 'normal' lane with aberrant transcript, subsequently shown to have tumour cells on cryosection). **d**) HSC73 RT-PCR of the cell lines, showing aberrant bands in MDA-MB435, MDA-MB468 and U937, but not in MDA-MB453 or T47D. **e**) Western blot of paired cases (**N/T**) run as RT-PCR samples above, probed with polyclonal antibody to HSC73 N-terminal peptides, show aberrant protein bands in 7 of the 9 tumour lanes. **f**) Western blot with the cell lines MDA-MB453 (**2**) and T47D (**4**) in addition to those above, shows a truncated HSC73 isoform in U937.

Figure 4- a) Genomic and mRNA structure of HSC73 showing the localisation of the probes (black bars) and PCR primers (arrows) used. The upper genomic structure has 9 exons (coding exons: stippled boxes; non-coding regions of exons: stippled/black boxes). The dashed introns (5, 6 and 8) contain the U14 SNURPS. Note the antibody epitope (=) and RNA-ISH probe (-) sites on HSC73 mRNA.

Figure 5- Human HSC73 gene and amino acid sequence (Y00371)

Figure 6 - Murine HSC73 cDNA and amino acid sequence (U27129)

Figure 7- Three BALB/c nude mice were injected subcutaneously with 500,000 cells from each cell line. After six weeks the mice were sacrificed and tumours evaluated.

Lane 1: mda435 parental line (tumour weights 8, 30 and 18mg)

Lane 2: mda435 HSC73 transfectant 1 (no tumours)

Lane 3: mda435 HSC73 transfectant 2 (one tumour of 68mg)

Lane 4: mda435 HSC73 transfectant 3 (no tumours)

Notes. The tumour which grew in one mouse injected with mda435 HSC73 transfectant 2 is thought to have arisen from a cell which had deleted the transfected clone. In the mice the cells are no longer under selection.

References

1. Stickland, J.E., Tomlinson, I.P.M., Lee, A.S.G., Evans, M.F. and McGee, J.O. (1992). Allele loss on chromosome 11q is a frequent event in breast cancer (abstract). *Br .J.Cancer*, **66,** (Suppl 17), 3.

2. Tomlinson, I.P., Stickland, J.E., Lee ASG, Bromley L, Evans M.F., Morton J. and McGee, J.O. (1995). *J. Clin Pathol.*, 48, 424-428

3. Koreth, J., Bethwaite, P.B. and McGee, J.O. (1995). Mutation at chromosome 11q23 in human nonfamilial breast cancer: a microdissection analysis. *J. Pathol.,* **176,** 11-18.

4. Koreth, J., Bakkenist, C.J. and McGee, J.O. (1997). Allelic deletions at chromosome 11q22-q23.1 and 11q25-qterm are frequent in sporadic breast but not colorectal cancers.*Oncogene,* **14,** 431-437

5. Negrini, M., Sabbioni, S., Possati, L., Rattan, S., Corallini, A., Barbanti, B.G. and Croce, C.M. (1994). Suppression of tumorigenicity of breast cancer cells by microcell-mediated chromosome transfer: studies on chromosomes 6 and 11. *Cancer Res.,* **54,** 1331-1336.

6. Hampton, G.M., Mannermaa, A., Winquist, R., Alavaikko, M., Blanco, G., Taskinen, P.J., Kiviniemi, H., Newsham, I., Cavenee, W.K. and Evans, G.A. (1994). Loss of heterozygosity in sporadic human breast carcinoma: a common region between 11q22 and 11q23.3. *Cancer Res.,* **54,** 4586-4589.

7. Bethwaite, P.B., Koreth, J., Herrington, C.S. and McGee, J.O. (1995). Loss of heterozygosity occurs at the D11S29 locus on chromosome 11q23 in invasive cervical carcinoma. *Br. J. Cancer,* **71,** 814-818.

8. Tomlinson, I.P., Gammack, A.J., Stickland, J.E., Mann, G.J., MacKie, R.M., Kefford, R.F. and McGee, J.O. (1993). Loss of heterozygosity in malignant melanoma at loci on chromosome 11 and 17 implicated in the pathogenesis of other cancers.*Genes Chrom. Cancer*, **7,** 169-172.

9. Evans, M.E., Herrington, C.S., Koreth,, J., Bakkenist, C.J. and McGee, J.O. (1997). Evidence for clonality and loss of heterozygosity at 11q23 in CIN III lesions. (manuscript in preparation).

10. Foulkes, W.D., Campbell, I.G., Stamp, G.W. and Trowsdale, J. (1993). Loss of heterozygosity and amplification on chromosome 11q in human ovarian cancer. *Br. J. Cancer,* **67,** 268-273.

11. Baffa, R., Negrini, M., Mandes, B., Rugge, M., Ranzani, G., Hirohashi, S. and Croce, C.M. (1996). Loss of heterozygosity for chromosome 11 in adenocarcinoma of the stomach. *Cancer Res.,* **56,** 268-273.

12. Rasio, D., Negrini, M., Manenti, G., Dragani, T.A. and Croce, C.M. (1995). Loss of heterozygosity at chromosome 11q in lung adenocarcinoma: identification of three independent regions. *Cancer Res.,* **55,** 3988-3991.

13. Koreth, J., Bakkenist, C.J. and McGee, J.O. (1997). Chromosome 11 and cancer: a review. (manuscript in preparation).

14. Trent J., Yang, J-M., Emerson, J., Dalton, W., McGee, D., Massey, K., Thompson, F. and Villar, H. (1993). Clonal chromosome abnormalities in human breast carcinomas. II. Thirty-four cases with metastatic disease. *Genes Chrom. Cancer*, **7,** 194-203.

15. Herrington, C.S., Leek, R.D. and McGee, J.O'D. (1995). Correlation of numerical chromosome 11 and 17 imbalance with metastasis of primary breast cancer to lymph nodes. *J. Pathol.* **176,** 353-359.

16. Saxon, P.J., Srivatsan, E.S. and Stanbridge, E.J. (1986). Introduction of human chromosome 11 via microcell transfer controls tumorigenic expression of HeLa cells. *EMBOJ.,* **5,** 3461-3466.

17. Parshad, R., Price, F.M., Oshimura, M., Barrett, J.C., Satoh, H., Weissman, B.E., Stanbridge, E.J. and Sanford, K.K. (1992). Complementation of a DNA repair deficiency in six human tumor cell lines by chromosome 11. *Hum . Genet.,* **88,** 524-528.

18. Robertson, G., Coleman, A. and Tracey, G.L. (1996). A malignant melanoma tumor suppressor on human chromosome 11. *Cancer Res.,* **56,** 4487-4490.

19. Phillips, K.K., Welch, D.R., Miele, M.E., Lee, J-H., Wei, L.L. and Weissman, B.E. (1996). Suppression of MDA-MB-435 breast carcinoma cell metastasis following the introduction of human chromosome 11. *Cancer Res.,* **56,** 1222-1227.

20. Dworniczak, B., Mirault, M.E. (1987). Structure and expression of a human gene coding for a 71kd heat shock 'cognate' protein. *Nuc. AcidsRes.* **15,** 5181-5197.

21. Tavaria, M., Gabriele, T., Anderson, R.L., Mirault, M-E, Baker, E., Sutherland, G. and Kola, I. (1995). Localization of the gene encoding the human heat shock cognate protein, HSP73, to chromosome 11. *Genomics* **29,** 266-268.

22. Brown, R.C., Hong-Brown, L.Q., Doxsey, S.J., Welch, W.J. (1996). Molecular Chaperones and the Centrosome. *J. Biol. Chem.* **271,** 833-840.

23. Brown, R.C., Martin, R.L., Hansen, W.J., Beckmann, R.P., Welch, W.J. (1993). The constitutive and stress inducible forms of hsp70 exhibit functional similarities and interact with one another in an ATP dependent fashion. *J. Cell Biol.* **120,** 1101-1112.

24. Bercovich, B., Stancovski, I., Mayer, A., Blumfield, N., Laszlo, A., Schwartz, A.L., Ciechanover, A. (1997). Ubiquitin-dependent

Degradation of Certain Protein Substrates in Vitro Requires the molecular Chaperone Hsc70. *J. Biol. Chem.* **272,** 9002-9010.

25. Brodsky, J.L. (1996). Post-translational protein translocation: not all hsc70s are created equal. *TIBS* **21,** 122-126.

26. Barbhaiya, H., Leverette, R., Liu, J., Maxwell, E.S. (1994). Processing of U14 small nucleolar RNA from three different introns of the mouse 70kDa-cognate-heat-shock-protein pre-messenger RNA. *Eur, J. Biochem.* **226,** 765-771.

27. Giebel, L.B., Dworniczak, B.P., and Bautz, E.K.F. (1988). Developmental regulation of a constitutively expressed mouse mRNA encoding 72kDa heat shock-like protein. *Dev. Biol.* **125**, 200-207.

28. Soulier, S., Vilotte, J-L., L'Huiller, J. and Mercier, J-C. (1996). Developmental regulation of murine integrin B1 subunit- and HSC73-encoding genes in mammary gland: sequence of a new mouse HSC73 cDNA. *Gene* **172**, 285-289.

29. Harris, A.L. (1992). p53 expression in human breast cancer. *Adv. Cancer Res.* **59**, 69-88.

30. Koonin, E.V. and Abagyan, R.A.. (1997). TSG101 may be the prototype of a class of dominant negative ubiquitin regulators. *Nat. Genet.* **16**, 330-31.

31. Haber, D. and Harlow, E. (1997). Tumour suppressor genes: evolving definitions in the genomic age. *Nat. Genet.* **16**, 320-22.

32. O'Malley, K., Mauron, A., Brachas, J.D. and Kedes, L. (1985). Constitutively expressed rat mRNA encoding a 70-kilodalton heat-shock-like protein. *Mol. Cell Biol.* **5**, 3476-83.

33. Trinh, R.Q. and Maxwell, E.S. (1988). Homologous genes for mouse 4.5S hybRNA are found in all eukaryotes and their low molecular weight RNA transcripts intermolecularly hybridize with eukaryotic 18S ribosomal RNAs. *Nucleic Acids Res.* **16**, 6041-56.

34. Sager, R. (1997). Expression genetics in cancer: shifting the focus from DNA to RNA. *Proc. Natl. Acad. Sci. U S A* **94**, 952-5.

35. Rice, L.W., Jazaeri, A.A. and Shupnik, M.A. (1997). Estrogen receptor mRNA splice variants in pre- and postmenopausal human endometrium and endometrial carcinoma. *Gynecol. Oncol.* **65**, 149-57.

36. Zhang, Q.X., Hilsenbeck, S.G., Fuqua, S.A. and Borg, A. (1996). Multiple splicing variants of the estrogen receptor are present in individual human breast tumors. *J. Steroid Biochem. Mol. Biol.* **59**, 251-60.

CLAIMS

1. A method of diagnosis performed on a biological sample, which method comprises detecting the aberrant expression of a gene at the HSC73 locus, or detecting a mutation at the HSC73 locus, in the sample.

2. The method according to claim 1, for diagnosing neoplasia or a susceptibility to neoplasia.

3. The method according to claim 1 or claim 2, wherein detection is at the protein level.

4. The method according to claim 3, wherein the amount of HSC73 protein is detected by means of an antibody specific for HSC73.

5. The method according to claim 4, wherein the antibody is specific for an epitope encoded by a region of the HSC73 gene which bridges exons 8 and 9.

6. The method according to claim 1 or claim 2, wherein detection is at the DNA or RNA level.

7. The method according to claim 6, wherein HSC73 DNA or RNA is detected by means of a specific probe.

8. The method according claim 6 or claim 7, wherein an abnormality in exon 2 or exon 3 is detected.

9. The method according to claim 8, wherein the abnormality is a deletion in intron 2.

10. The method according to claim 8, wherein the abnormality is a point mutation in exon 3.

11. The method according to any one of claims 1 to 10, wherein the biological sample is derived from breast, ovary, prostate or lung tissue.

12. A DNA which encodes a mutated or unusual polymorphic form of an HSC73 mRNA or polypeptide, fragments of the DNA including the mutation or polymorphism and complementary DNAs.

13. The DNA according to claim 12, containing an abnormality in at least one of intron 2 and exon 3.

14. An RNA molecule encoded by the DNA according to claim 12 or claim 13.

15. A polypeptide encoded by the DNA according to claim 12 or claim 13.

16. Oligonucleotides capable of specifically hybridising with a region of an HSC73 DNA or RNA according to any one of claims 12 to 14, which region includes the mutation or polymorphism.

17. The oligonucleotides according to claim 16, labelled with a detectable moiety.

18. Antibodies which bind specifically to a region of an HSC73 protein encoded by exon 3.

19. The antibodies according to claim 18, wherein the exon 3 is a mutated or unusual polymorphic form of HSC73 exon 3.

20. Synthetic peptides for use in raising antibodies according to claim 18 or claim 19.

21. A genetically modified non-human mammal having genetic weakness to disease, in which the somatic cells and germ cells have at least one mutated or deleted copy of HSCC73.

22. An embryonic stem cell transformed with a nucleic acid capable of replacing a copy of HSC73 in the cell.

23. A cell line genetically engineered to have at least one mutated or deleted copy of HSC73.

24. A kit comprising an oligonucleotide according to claims 16 or 17 or an antibody according to claims 18 or 19.

25. A composition comprising a gene delivery vector which comprises a nucleic acid encoding an HSC73 protein under the control of a suitable promoter, together with a pharmaceutically acceptable carrier.

26. A method of treatment comprising introducing into a patient at least one copy of a nucleic acid which encodes a native HSC73 under the control of a suitable promoter.

Fig. 1.

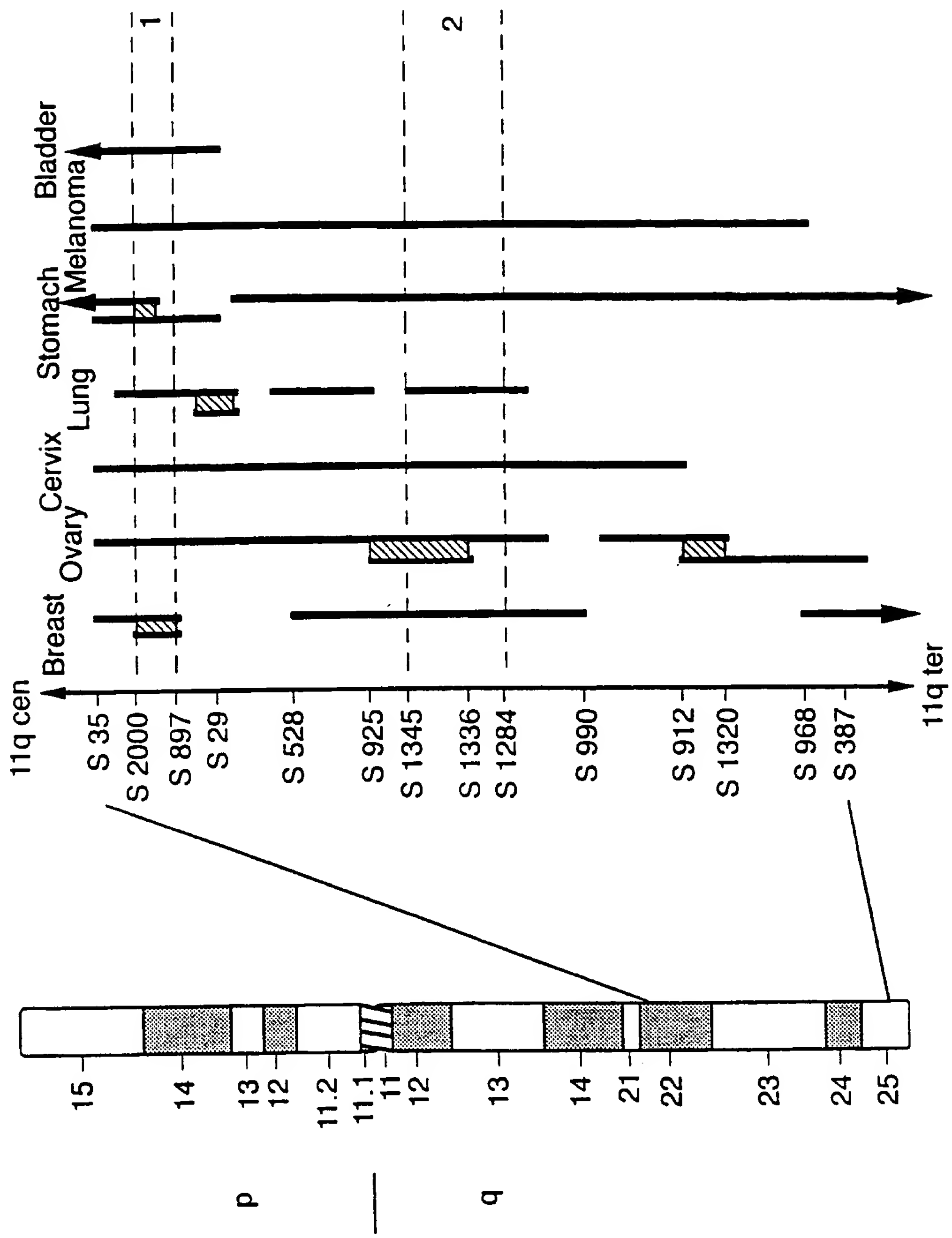

Breast
Ovary
Cervix
Lung
Stomach
Melanoma
Bladder
1
2
11q cen
S 35
S 2000
S 897
S 29
S 528
S 925
S 1345
S 1336
S 1284
S 990
S 912
S 1320
S 968
S 387
11q ter
15
14
13
12
11.2
11.1
11
12
13
14
21
22
23
24
25
p
q

Fig.2.

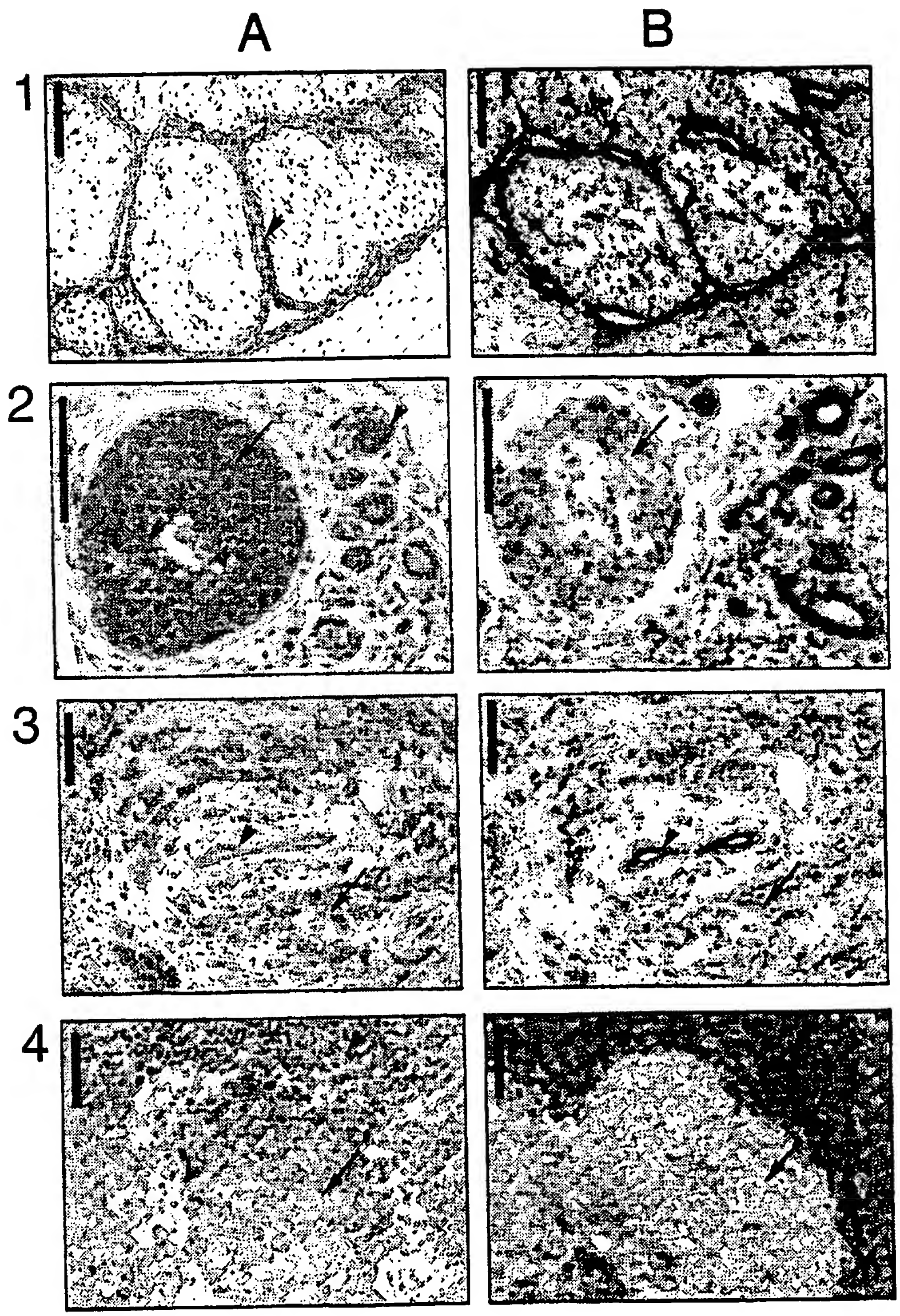


Fig.3.

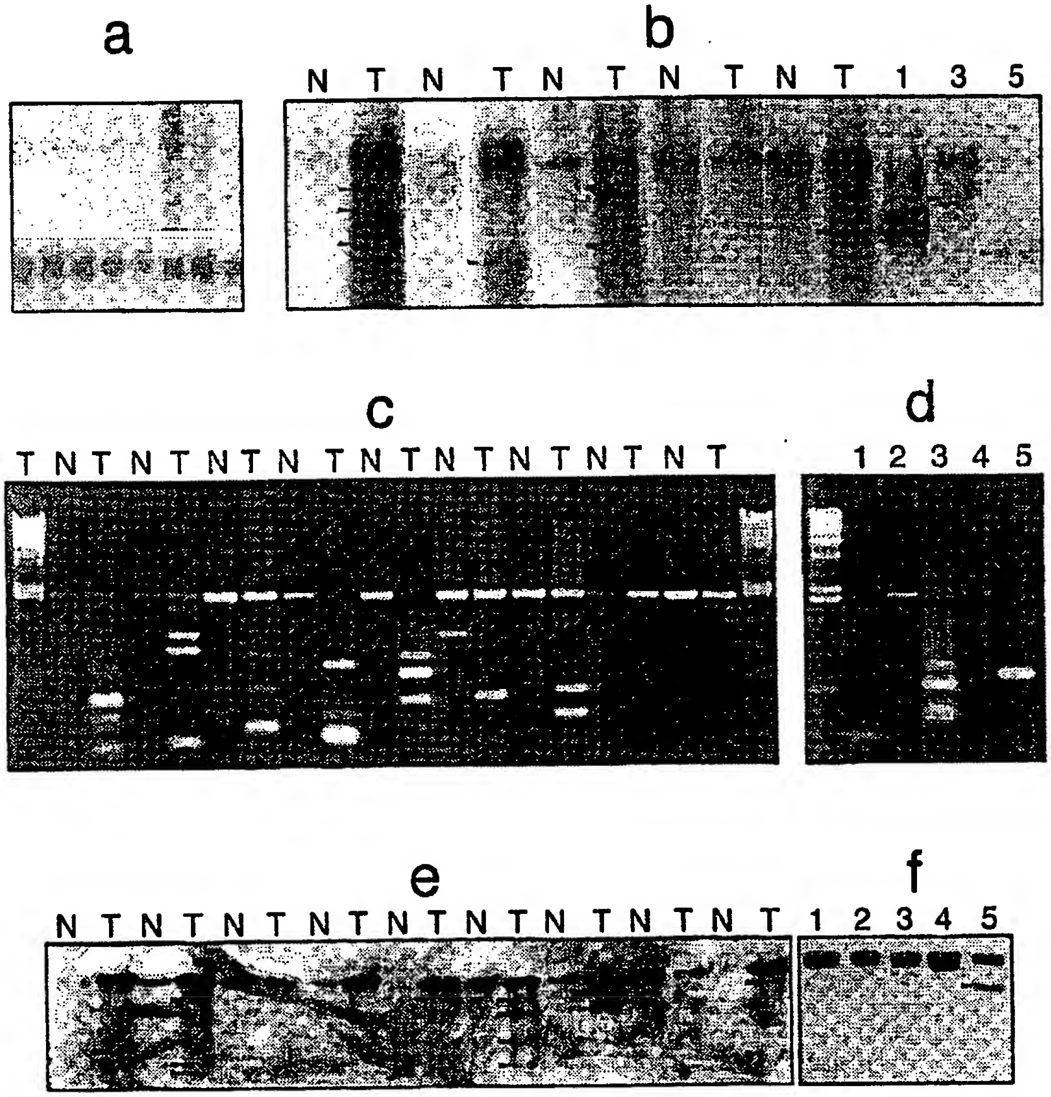

a
b
N T N T N T N T N T 1 3 5
c
T N T N T N T N T N T N T N T N T N T
d
1 2 3 4 5
e
N T N T N T N T N T N T N T N T N T
f
1 2 3 4 5

Fig.4.

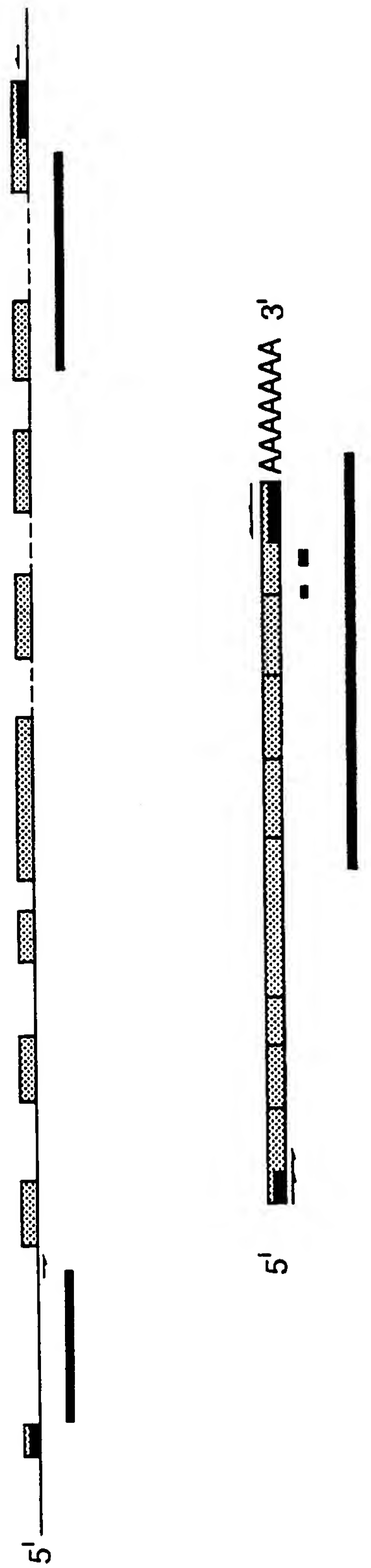


Figure 5

```
Key               Location/Qualifiers

source            1..5408
                  /organism="Homo sapiens"
                  /clone="phage521"
                  /clone_lib="human liver DNA in lambda L47.1"
misc_feature      4..17
                  /note="heat shock consensus element"
misc_feature      109..113
                  /note="regulatory sequence"
misc_feature      178..191
                  /note="heat shock consensus element"
TATA_signal       199..203
prim_transcript   227..>4630
mRNA              join(227..304,1035..1244,1567..1772,2097..2249,2337..2892,
                  3104..3306,3535..3733,3881..4113,4445..>4630)
exon              227..304
                  /number=1
intron            305..1034
                  /number=1
exon              1035..1244
                  /number=2
CDS               join(1040..1244,1567..1772,2097..2249,2337..2892,
                  3104..3306,3535..3733,3881..4113,4445..4630)
                  /db_xref="PID:g32467"
                  /db_xref="SWISS-PROT:P11142"
                  /product="71 Kd heat shock cognate protein"
                  /translation="MSKGPAVGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFT
                  DTERLIGDAAKNQVAMNPTNTVFDAKRLIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQ
                  VEYKGETKSFYPEEVSSMVLTKMKEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGT
                  IAGLNVLRIINEPTAAAIAYGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKST
                  AGDTHLGGEDFDNRMVNHFIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQAS
                  IEIDSLYEGIDFYTSITRARFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGS
                  TRIPKIQKLLQDFFNGKELNKSINPDEAVAYGAAVQAAILSGDKSENVQDLLLLDVTPL
                  SLGIETAGGVMTVLIKRNTTIPTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLGK
                  FELTGIPPAPRGVPQIEVTFDIDANGILNVSAVDKSTGKENKITITNDKGRLSKEDIER
                  MVQEAEKYKAEDEKQRDKVSSKNSLESYAFNMKATVEDEKLQGKINDEDKQKILDKCNE
                  IINWLDKNQTAEKEEFEHQQKELEKVCNPIITKLYQSAGGMPGGMPGGFPGGGAPPSGG
                  ASSGPTIEEVD"
intron            1245..1566
                  /number=2
exon              1567..1772
                  /number=3
intron            1773..2096
                  /number=3
exon              2097..2249
                  /number=4
intron            2250..2336
                  /number=4
exon              2337..2892
                  /number=5
intron            2893..3103
                  /number=5
exon              3104..3306
                  /number=6
```

intron	3307..3534	
	/number=6	
exon	3535..3733	
	/number=7	
intron	3734..3880	
	/number=7	
exon	3881..4113	
	/number=8	
intron	4114..4444	
	/number=8	
exon	4445..>4630	
	/number=9	
polyA_signal	4843..4848	

Sequence 5408 BP; 1411 A; 1079 C; 1420 G; 1498 T; 0 other;
GAGCTTGAAA GTTCCAGAAC GCTGCGGTGA GTGCGTTATC GTGAGGCGGC GCGGTGGGGT 60
GGGTGCGGAA GGGGGCGAGG CGAGGAGTGG AGCCGCGTTG TGATTGTGAT TGGGTCTTGT 120
AAGGGCAGCC GGACTCTATT GGCCGGGAAC CTAATGCAGG AAGCAGGCGG ACCCCTTCTG 180
GAAGGTTCTA AGATAGGGTA TAAGAGGCAG GGTGGCGGGC GGAAACCGGT GCTCAGTTGA 240
ACTGCGCTGC AGCTCTTGGT TTTTTGTGGC TTCCTTCGTT ATTGGAGCCA GGCCTACACC 300
CCAGGTAAAA CCTCTGCTCA AGAGTTGGGT TGTGGGTCTG GGAGCGTGCA GCCTCCACAC 360
AGGCCTGTTG GGCTTGCTGA GGCTTGGGGG TTCTGAGAAT CTCGTCGAGG CGAGTGTGCG 420
GCTCCTTCTA CCGGCTTAAA GGGCCTCAGT TTTCGGTGGG ATGGCAGCGG TATTTGGTTG 480
CAGCCGGCAG ACGGAAATGT AGGGAGTGGG CCGCATGGCC CCAGGGGAGG CTGGGAGACG 540
CCCGGCCGCG TGGCGGGGGA GGGTTGCTGC ATCGGTTTGC CTGGCGCGCG GGGAAGTGGA 600
GCCAGCGTTT TCTTTCACCC AGTTCCCTGC TTAGTCCAGT CCCACCGTGG TTCTTCAGAG 660
CTGTTCTTGG CGTGCTTCCA GTATGGGGGT ACATTCCGGA GTAGTTAAAA GCCCGTTGAC 720
TCCCGGGGGG CACTGGCACC TGGCGAGGGA GGGGAACAGA CAGTGCTCAG TTCGGGGTAA 780
GACCACGTGT TGAGCAACGC CCCACGCCGT CTGGGTCGAT GGGTCCTTCA TCTAGGGCGT 840
GCTGTGCTGC GGTTGGCACG GCAACCTGGA CTGCAGCACT AGTTCTGGAC CTCGCGCGTG 900
CTTAGACAGG AGGTGATGGG CACTATTACC TCTTGGCAGT GGCCATACGT TTTTCCTGGT 960
TAAGTGTTCT GTTAAGGGAT GAGGGAAATA TTTTGATTAA TTGAATTTTT AAACCAGATT 1020
TTTCTTTTTT TCAGCAACCA TGTCCAAGGG ACCTGCAGTT GGTATTGATC TTGGCACCAC 1080
CTACTCTTGT GTGGGTGTTT TCCAGCACGG AAAAGTCGAG ATAATTGCCA ATGATCAGGG 1140
AAACCGAACC ACTCCAAGCT ATGTCGCCTT TACGGACACT GAACGGTTGA TCGGTGATGC 1200
CGCAAAGAAT CAAGTTGCAA TGAACCCCAC CAACACAGTT TTTGGTGAGT TCCTAATTTT 1260
AAATGACAGA ACAAATATAA ACAGGGCTAG GAAGCACAAA AGTTTATGAA ACGTGAGGAG 1320
GGAACTTTTT GATTTTAGAA AAACTGAGCT GAGAGACTTG TTATCAAGTC TGTTATAAAA 1380
CAGGTTGTAG AAACCTTTCA GGCTGAAATC TGGATAACGT AGGAGGTTGA AGTTTGAACC 1440
TTTGCTAGGT ATATGGTAGT TGAATTCACC TACCTATGAA CTGTTAGGTA TTTGAGTAAT 1500
CATGGACTTG AGTTTTATCT GAAGAGCTAT GAAATTGAAA GTGTTTTCAT TTGACACCTT 1560
TTACAGATGC CAAACGTCTG ATTGGACGCA GATTTGATGA TGCTGTTGTC CAGTCTGATA 1620
TGAAACATTG GCCCTTTATG GTGGTGAATG ATGCTGGCAG GCCCAAGGTC CAAGTAGAAT 1680
ACAAGGGAGA GACCAAAAGC TTCTATCCAG AGGAGGTGTC TTCTATGGTT CTGACAAAGA 1740
TGAAGGAAAT TGCAGAAGCC TACCTTGGGA AGGTGAGGTT GGTTTTTCAG TATGGGGTGC 1800
ATTCCGGAGT AGTTAAAAGC CCGATGACTC CCGGGGGCAC TGGCACCTGG CGAGGGAGGG 1860
GAACAGATGG GGCTCAGCTC AGGGTTAAGA CCACGTGCCC AACAGTGCCC TAGGCTCTCT 1920
AGGTAGATGG GTCTGTCAAC ACCAGAAACC AGTGAATCTT GACAATTACA CAGTAATTTA 1980
CATTTTGGTG GGGGGGGTGC TCCAGCTGTT GTTTCACCAG CATTAATCCA TTTGCTGGAG 2040
TTTGCATATA TGTAAGTATA ATAGTTACCA ATCTGTGGTC TTTTCCTTAT TCCTAGACTG 2100
TTACCAATGC TGTGGTCACA GTGCCAGCTT ACTTTAATGA CTCTCAGCGT CAGGCTACCA 2160
AAGATGCTGG AACTATTGCT GGTCTCAATG TACTTAGAAT TATTAATGAG CCAACTGCTG 2220
CTGCTATTGC TTACGGCTTA GACAAAAAGG TATGTACCAT TTGTGATGCA AGTTCGGATT 2280
ATTTTAAGAT TAATTTGATC CATCGTAAAT TTAAATGAGA TTGTTTTTAA CGGCAGGTTG 2340
GAGCAGAAAG AAACGTGCTC ATCTTTGACC TGGGAGGTGG CACTTTTGAT GTGTCAATCC 2400
TCACTATTGA GGATGGAATC TTTGAGGTCA AGTCTACAGC TGGAGACACC CACTTGGGTG 2460
GAGAAGATTT TGACAACCGA ATGGTCAACC ATTTTATTGC TGAGTTTAAG CGCAAGCATA 2520
AGAAGGACAT CAGTGAGAAC AAGAGAGCTG TAAGACGCCT CCGTACTGCT TGTGAACGTG 2580

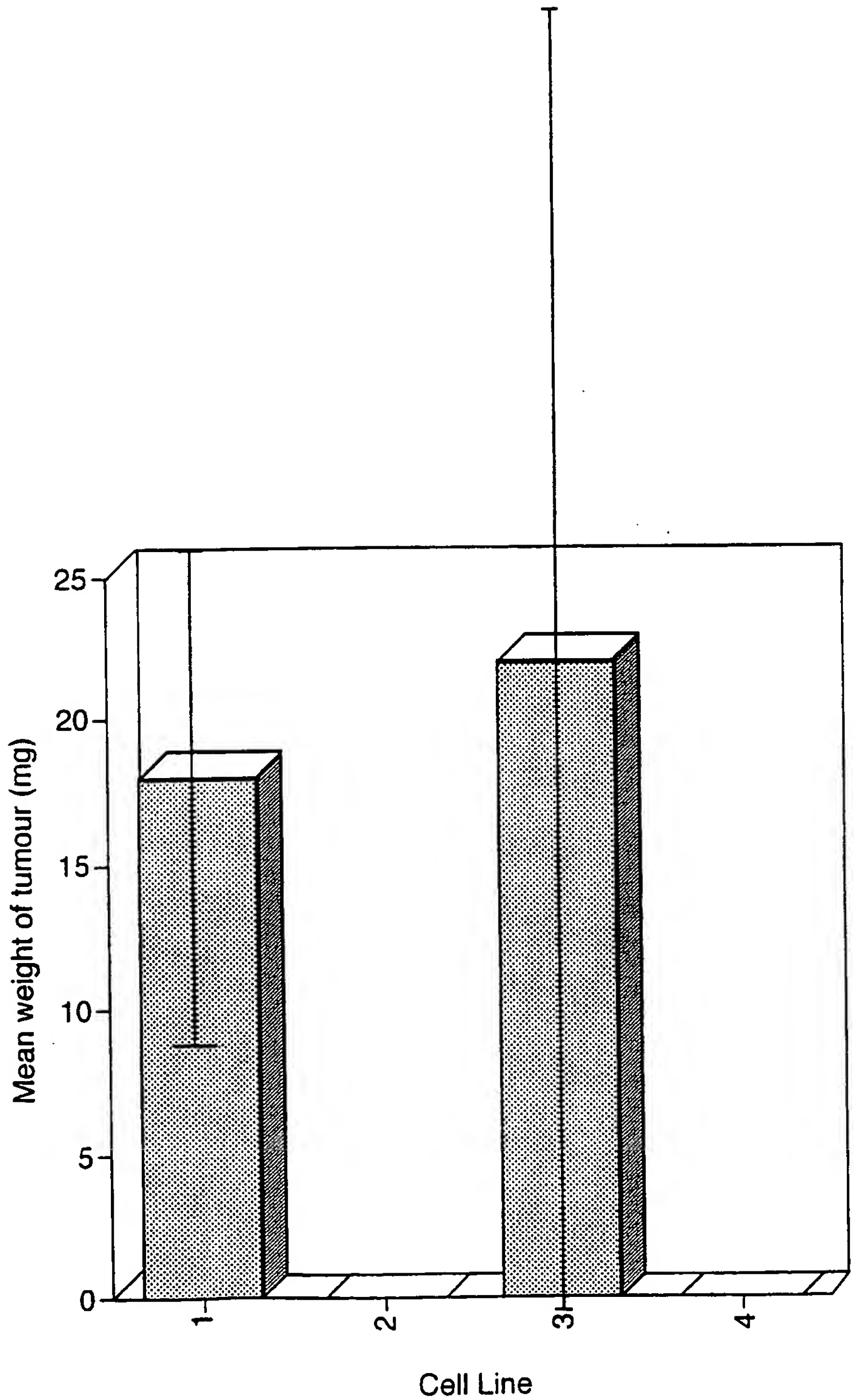
```
CTAAGCGTAC CCTCTCTTCC AGCACCCAGG CCAGTATTGA GATCGATTCT CTCTATGAAG 2640
GAATCGACTT CTATACCTCC ATTACCCGTG CCCGATTTGA AGAACTGAAT GCTGACCTGT 2700
TCCGTGGCAC CCTGGACCCA GTAGAGAAAG CCCTTCGAGA TGCCAAACTA GACAAGTCAC 2760
AGATTCATGA TATTGTCCTG GTTGGTGGTT CTACTCGTAT CCCCAAGATT CAGAAGCTTC 2820
TCCAAGACTT CTTCAATGGA AAAGAACTGA ATAAGAGCAT CAACCCTGAT GAAGCTGTTG 2880
CTTATGGTGC AGGTAACAAT GGTATCTCAA TTAACCCTAA AGGCAGGCAG GCCCAAGGTG 2940
ACTCGCTGTG ATGAGTGATT GTTAAACATT CGTAGTTTCC ACCAAAAGCT TGGCTAATGA 3000
TGGCAACACC TTCCTTGGAT GTCTGAGCGA GTGATAGTTA AAACAGGAGC TATGTACTGG 3060
GTTTTCTTTT AACTTCTTTT AACGTTAACT TTTTGTTTGC TAGCTGTCCA GGCAGCCATC 3120
TTGTCTGGAG ACAAGTCTGA GAATGTTCAA GATTTGCTGC TCTTGGATGT CACTCCTCTT 3180
TCCCTTGGTA TTGAAACTGC TGGTGGAGTC ATGACTGTCC TCATCAAGCG TAATACCACC 3240
ATTCCTACCA AGCAGACACA GACCTTCACT ACCTATTCTG ACAACCAGCC TGGTGTGCTT 3300
ATTCAGGTAT GTTTCTGTAC TTCTCTTGTT TGGCTTACTG ATAACAGATA AAGGGAAGTC 3360
TTGACTGACT CGCTATGATG ATGGATTCCA AAACCATTCG TAGTTTCCAC CAGAAAGTCT 3420
TATGTTGGCC AGTTCCTTCC TTGGATGTTT GAGCGACCAT TCTTCCTTAG CAGGACCCTA 3480
GCACTGTCAC AGACCTGGAG TCCATTGTAG TAATTTGTTT TATTTCCTAC CAAGGTTTAT 3540
GAAGGCGAGC GTGCCATGAC AAAGGATAAC AACCTGCTTG GCAAGTTTGA ACTCACAGGC 3600
ATACCTCCTG CACCCCGAGG TGTTCCTCAG ATTGAAGTCA CTTTTGACAT TGATGCCAAT 3660
GGTATACTCA ATGTCTCTGC TGTGGACAAG AGTACGGGAA AAGAGAACAA GATTACTATC 3720
ACTAATGACA AGGGTAAGGA GGCACTGTCA TCTGGTCTTG ACAGGGATAA TGGTATTTCA 3780
ATTGAGTTAC TGGTGAATAA GGGCGTCTAG CTAAGAGAAA CTAGAGTTAC ACATACACAG 3840
GTAATTTAAG GCTTTTACTT AGAGTTAATT TCTTTCCTAG GCCGTTTGAG CAAGGAAGAC 3900
ATTGAACGTA TGGTCCAGGA AGCTGAGAAG TACAAAGCTG AAGATGAGAA GCAGAGGGAC 3960
AAGGTGTCAT CCAAGAATTC ACTTGAGTCC TATGCCTTCA ACATGAAAGC AACTGTTGAA 4020
GATGAGAAAC TTCAAGGCAA GATTAACGAT GAGGACAAAC AGAAGATTCT GGACAAGTGT 4080
AATGAAATTA TCAACTGGCT TGATAAGAAT CAGGTTTGTG TTTTTTTTTT TTTTTTTCCT 4140
CCCCCACGCA ATGGAGGGGA AGGGGATGGT AAACCAAGCT TGAGCTGGAT TTCAGTGTAG 4200
GGTCACAATG ATGAATGGTC CAAAACATTC GCGGTTTCCA CCAGAATTCA AGGTGTTGGC 4260
AACTACCTTC CTTGGATGTC TGAGTGACCC AAGATGTTAA GGAAGAATAA GGCCCTATTT 4320
TAATGTTGGT ATGGGCCCTC TTGTAAGAGT TTGCTCCAGA CTTTTAGTAT CAGATTGCGT 4380
CAGGGAGAAA GAAGGGTTAT TAACATTAAA AGAACTTGCA GTAATTCCTT TTTCTCTTCC 4440
TCAGACTGCT GAGAAGGAAG AATTTGAACA TCAACAGAAA GAGCTGGAGA AAGTTTGCAA 4500
CCCCATCATC ACCAAGCTGT ACCAGAGTGC AGGAGGCATG CCAGGAGGAA TGCCTGGGGG 4560
ATTTCCTGGT GGTGGAGCTC CTCCCTCTGG TGGTGCTTCC TCAGGGCCCA CCATTGAAGA 4620
GGTTGATTAA GCCAACCAAG TGTAGATGTA GCATTGTTCC ACACATTTAA AACATTTGAA 4680
GGACCTAAAT TCGTAGCAAA TTCTGTGGCA GTTTTAAAAA GTTAAGCTGC TATAGTAAGT 4740
TACTGGGCAT TCTCAATACT TGAATATGGA ACATATGCAC AGGGGAAGGA AATAACATTG 4800
CACTTTATAC ACTGTATTGT AAGTGGAAAA TGCAATGTCT TAAATAAAAC TATTTAAAAT 4860
TGGCACCATA CAATTGCTTT GAGTCTTTAA ATAATCTCCC AGGCCAGCGG TGGGAGAAGT 4920
AGGCTTAGGT GATTATGTGA CTCTTACTTT CTCCTTCCTC TTAAGCTTGA GTTAACAAGG 4980
GCTGGGTGGC AAGTTGCCCT TCAGAGCATG TGGATGGTAC ATTTTGGAAT TCAGAGCTTT 5040
GAGAAGGGGA GCATAAGAAA TTGGATCTGG ATCAAACTAA CCTTAGTCCT TAGGCTGGAG 5100
AGGCAGAAGC TGACTTAATG GTGTTTTCTA AACTTATTCT GTGTGTAAGC CTGCCTAGGA 5160
GCAGAGGCTT TCCTGGAGGG TTGTGCTAGA TGAGTAAGAA TTTAGATACA GAATCAAATA 5220
ATGGGCAGTG AATATTAAGC TACATGGCAG AGGTATCTGA ATGTCAATCC CTTATATGAG 5280
CCACTGCCCT GTGGGCTTCC ATTTCTTCTG AGTTAAGATT ATTCAGAAGG TCGGGGATTG 5340
GAGCTAAGCT GCCACCTGGT TAATTAAGGT CCCAACAGTG AGTTGTGATA GCCTAGGGGA 5400
GCAGGCTG 5408
```

Figure 6

```
FEATURES             Location/Qualifiers
  gene               31..1971
                     /gene="hsc73"
  CDS                31..1971
                     /gene="hsc73"
                     /codon_start=1
                     /product="heat shock 73 protein"
                     /db_xref="PID:g861213"
                     /translation="MSKGPAVGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAF
                     TDTERLIGDAAKNQVAMNPTNTVFDAKRLIGRRFDDAVVQSDMKHWPFMVVNDAGRPK
                     VQVEYKGETKSFYPEEVSSMVLTKMKEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKD
                     AGTIAGLNVLRIINEPTAAAIAYGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFE
                     VKSTAGDTHLGGEDFDNRMVNHFIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSS
                     STQASIEIDSLYEGIDFYTSITRARFEELNADLFRGTLDPVEKALRDAKLDKSQIHDI
                     VLVGGSTRIPKIQKLLQDFFNGKELNKSINPDEAVAYGAAVQAAILSGDKSENVQDLL
                     LLDVTPLSLGIETAGGVMTVLIKRNTTIPTKQTQTFTTYSDNQPGVLIQVYEGERAMT
                     KDNNLLGKFELTGIPPAPRGVPQIEVTFDIDANGILNVSAVDKSTGKENKITITNDKG
                     RLSKEDIERMVQEAEKYKAEDEKQRDKVSSKNSLESYAFNMKATVEDEKLQGKINDED
                     KQKILDKCNEIISWLDKNQTAEKEEFEHQQKELEKVCNPIITKLYQSAGGMPGGMPGG
                     FPGGGAPPSGGASSGPTIEEVD"

   1 cagcacagct gggcctacac acaagcaacc atgtctaagg gacctgcagt tggcattgat
  61 ctcggcacca cctactcctg tgtgggtgtc ttccagcatg gaaaggtgga aattattgcc
 121 aatgaccagg gtaaccgcac cacaccaagc tatgttgctt tcacggacac agagagatta
 181 attggggatg cggccaagaa tcaggttgca atgaacccca ccaacacagt tttgatgcc
 241 aaacgtctga tcgggcgtag gtttgatgat gctgttgttc agtctgatat gaagcactgg
 301 cccttcatgg tggtgaatga tgcaggcagg cccaaggtcc aagtggaata caaaggggag
 361 acaaaaagtt tctacccaga ggaagtgtcc tccatggttc tgacaaagat gaaggaaatt
 421 gcagaagcgt acctcggaaa gactgttacc aacgctgtgg tcacagtgcc cgcttacttc
 481 aatgactctc agcgacaggc aacaaaagat gctggaacta ttgctggcct caatgtactt
 541 cgaatcatca atgaaccaac tgctgctgct attgcttatg gcttagataa gaaggtcgga
 601 gctgaaagaa atgtgctcat ttttgacttg ggaggtggca cttttgatgt gtcaatcctc
 661 actattgagg atggaatttt tgaggtcaaa tcaacagctg gagacaccca cttaggtgga
 721 gaagattttg acaaccgaat ggtcaatcat ttcattgctg agttcaagcg aaagcacaag
 781 aaagacatca gtgagaacaa gagagctgtc cgccgtctcc gcacggcctg cgagcgggcc
 841 aagcgcaccc tctcctccag cacccaggcc agtattgaga ttgattctct ctatgaggga
 901 attgacttct atacctccat tacccgggct cgatttgagg agttgaatgc tgacctgttc
 961 cgtggcacac tggaccctgt agagaaggcc cttcgagatg ccaagctgga caagtcacag
1021 atccatgata ttgtcttggt gggtggttct accagaatcc ccaagattca gaaacttctg
1081 caagacttct tcaatggaaa agagctgaac aagagcatta accccgatga agctgttgcc
1141 tatggtgcag ctgtccaggc agccattcta tctggagaca agtctgagaa cgttcaggat
1201 ttgctgctct tggatgtcac tcctctttcc cttggtattg aaactgctgg cggagtcatg
1261 actgtcctca tcaagcgcaa taccaccatc cccaccaagc agacacagac tttcaccacc
1321 tactctgaca accagcctgg tgtactcatt caggtgtatg aaggtgaaag ggccatgacc
1381 aaggacaaca acctgcttgg aaagttcgag ctcacaggca tccctccagc accccgtggg
1441 gtccctcaga ttgaggttac ttttgacatc gatgccaatg gcatcctcaa tgtttctgct
1501 gtagataaga gcacaggaaa ggagaacaag atcaccatca ccaatgacaa gggccgcttg
1561 agtaaggaag atattgagcg catggtccaa gaagctgaga agtacaaggc tgaggatgag
1621 aagcagagag ataaggtttc ctccaagaac tcactggagt cctatgcctt caacatgaaa
1681 gcaactgtgg aagatgagaa acttcaaggc aagatcaatg atgaggacaa acagaagatt
1741 cttgacaagt gcaatgaaat catcagctgg ctggataaga accagactgc agagaaggaa
1801 gaatttgagc atcagcagaa agaactggag aaagtctgca accctatcat taccaagctg
1861 taccagagtg caggtggcat gcctggggga atgcctggtg gcttcccagg tggaggagct
1921 cccccatctg gtggtgcttc ttcaggcccc accattgaag aggtggatta agtcagtcca
1981 agaagaaggt gtagctttgt tccacaggga cccaaaacaa gtaacatgga ataataaaac
2041 tattcaaatt gg
```

Fig.7.



INTERNATIONAL SEARCH REPORT

Intern. ...al Application No
PCT/GB 98/03375

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68 C07K14/47 C07K16/18 C12N15/12 C12N15/00
C12N5/10 A61K38/16 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	UDONO H ET AL: "COMPARISON OF TUMOR-SPECIFIC IMMUNOGENICITIES OF STRESS-INDUCED PROTEINS GP96, HSP90, AND HSP70" JOURNAL OF IMMUNOLOGY, vol. 152, 1994, pages 5398-5403, XP002912526 see page 5398, paragraph 2 - page 5399, paragraph 1 ---	1,2,18
A	MULTHOFF G ET AL: "A stress-inducible 72-kDa heat-shock protein (hsp72) is expressed on the surface of human tumor cells, but not on normal cells" INTERNATIONAL JOURNAL OF CANCER, vol. 61, 10 April 1995, pages 272-279, XP002090859 see the whole document ---	1-20

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 March 1999	13/04/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Osborne, H

1

INTERNATIONAL SEARCH REPORT

Intern. .nal Application No
PCT/GB 98/03375

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	YANO M ET AL: "Expression and roles of heat shock proteins in human breast cancer" JAPANESE JOURNAL OF CANCER RESEARCH, vol. 89, no. 9, September 1996, pages 908-15, XP002098505 see the whole document ---	1-20
X	WO 97 06685 A (SLOAN KETTERING INST CANCER ;HOE MEE H (US); HOUGHTON ALAN (US); M) 27 February 1997	25,26
A	see the whole document ---	1,20-24
X	WO 94 11513 A (LUKACS KATALIN VERONIKA ;MEDICAL RES COUNCIL (GB); LOWRIE DOUGLAS) 26 May 1994	25,26
A	see the whole document ---	21-24
X	WO 96 10411 A (SINAI SCHOOL MEDICINE) 11 April 1996	25,26
A	see the whole document -----	21-24

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 98/03375

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9706685 A	27-02-1997	AU 6849396 A	12-03-1997
		AU 6898496 A	12-03-1997
		AU 6952896 A	12-03-1997
		CA 2229543 A	27-02-1997
		CA 2229595 A	27-02-1997
		EP 0851764 A	08-07-1998
		EP 0888053 A	07-01-1999
		WO 9706828 A	27-02-1997
		WO 9706821 A	27-02-1997
WO 9411513 A	26-05-1994	AU 690085 B	23-04-1998
		AU 5430794 A	08-06-1994
		CA 2149309 A	26-05-1994
		EP 0670899 A	13-09-1995
		JP 8503132 T	09-04-1996
WO 9610411 A	11-04-1996	US 5750119 A	12-05-1998
		AU 2281995 A	26-04-1996
		CA 2201498 A	11-04-1996
		CN 1167440 A	10-12-1997
		EP 0784477 A	23-07-1997
		JP 10506628 T	30-06-1998